**Informatics Unit 3**

**SAT – Criteria 4**

**Organising data**

**9-10 “**Thoroughly and systematically prepares the data for manipulation by selected software tools including codifying all qualitative data using appropriate **physical** techniques or software tools”.

* Prepares data for manipulation including correct data types and structures
* Digitise non-digital data.
* Codifying, qualitative data into categories using software; **physical** element eg. marking up a list of responses to a different question with different highlighters, or assigning a key to them, then tallying up the numbers.
* Code the data by grouping into like areas (this is the physical coding that is described in the performance descriptors)
* Once the data is coded, you use software tools, eg. Microsoft Excel, to develop a chart that can explain the relationship between the coded items in a more visually appealing way.

**Explain how you have met the above requirements.**

**Eg.**

*All the primary data collected will undergo data integrity test as follows:*

*Timeliness - This will be accomplished by completing my survey respondents at the same time*

*Authenticity - All the data collected will have email addresses of people surveyed and no one except me access to change or manipulate the survey or its results. Only people who have a school email address and password will be able to take this survey. No multiple copies of any data will be maintained in devices such as USB, flash drives or external hard disk drives.*

*Relevance - ………………..*

*Accuracy - The accuracy of the data collection method is quite sound. I am automatically collecting the email addresses from people who will be surveyed as long as they are logged onto the school network. The survey has been checked for completeness and questions have been locked to take only one response even though they click multiple responses.*

*Primary Data( Quantitative)*

*Total data surveyed =*

**Primary data source**

* *Two Google Forms created: Survey(quantitative) and Questionnaire (qualitative)*
* *Both forms are stored in relevant folders on the cloud*
* *Data collected is transferred into Google Sheets( download responses)*
* *Raw data is identified with unique file name ( original file)*
* *Tabs are created and data is copied for manipulation( tab names will be clearly identified)*
* *As Google drive is used all data is automatically backed by the service provider*

*Codifying………….*

**Secondary data source**

* This resource is downloaded from the below hyperlink:
* The raw .xls file saved in secured location on Google Drive
* The data selected will be ……………………..
* As Google drive is used all data is automatically backed by the service provider

**Manipulation of data**

9-10: Requirements regarding manipulation of data: Use extensive range of software, functions, techniques, formats & conventions to manipulate data to

* Generate insightful information that is suitable for interpretation, using appropriate software.
* Use of an extensive range of:
* digital system components,
* software functions,
* techniques,
* formats & conventions

**Explain how you have met the above requirements.**

### Eg. *Hardware used*

* *Personal Computer Quad Core or higher*
* *Windows 7 or higher (operating system)*
* *Minimum 2 GB RAM*
* *Minimum 128 GB Hardware storage*

### *Software Tools used*

* *Good drive installed on local PC*
* *Microsoft Excel and word*
* *Google Office Applications( Docs, Sheets, Forms)*

### *Backup Devices or services*

* *Cloud services( Google Drive)*

### Data manipulation

**Primary data**

* Responses are viewed in Google sheets and all 20 responses are downloaded as .xls on the desktop( within Google Drive application for automatic synchronization)
* The .xls file is now opened using Microsoft Office and examined.
* Fields that are not in use will be discarded ( such as ….)
* All necessary fields ( insert here…) is saved .xls files with appropriate column headers and saved. No backup is necessary due to automatic sync and backup.
* *Insert rules used and Statistical*  …..

**Storing & securing data and information and meeting legal requirements**

* Explain how the files have been secured using software security controls to protect all stored and communicated data & information.
* Explain how your files meet all the legal requirements (refer to relevant legislation).

Eg.

*No names or email address will be made available to the public to be ensure Privacy Act, Privacy and Data protection Act, Health Record Act (HRA) is strictly followed. Both data data sets collected will respect all Privacy requirements and will not identify any people taking the survey.*

*The original data sources will be completely protected to ensure Privacy of the participants are maintained. The only data available or published will be responses (without identification) in an .xls file. The secondary source of data (from ………) has a detailed online privacy policy that protects their data. No health record or medical information are collected to ensure the HRA is maintained.*

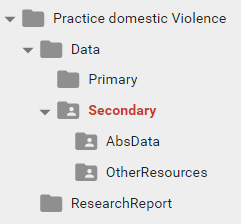
*no original data or secondary data will be made publicly or available online and appropriate referencing( APA style) will be completed.*

**File Management**

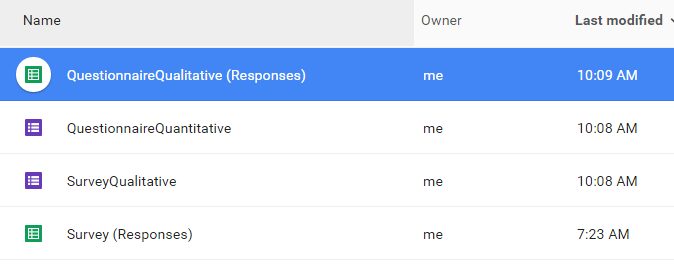
* Documented file management plan including a diagram of your folder structure.

### Data storage and Hierarchies

* Folder naming convention and organisation
* Simple and plain



* File naming convention ( Data⇒ Primary)
* CamelCase, Short names, plain, versions automatically managed by Google Drive
* Force logical order with “nameYYYY-MM-DD”



### Data file management

***Primary source of data***

1. *SurveyQuantitative(Responses) ⇒ Google sheets ( Original file generated from Google form survey).*
2. *QuestionnaireQualititative(Responses) ⇒ Google sheets ( Original file generated from Google form survey)*
3. *Google sheets ⇒ .xls. The raw data imported into .xls file*
4. *The raw data is selected and manipulated, sorted and filtered*
5. *Data processed and graphs and visualisation produced*
6. *Find document ( partially completed Google shared link ………………………………*

**Secondary source of data**

1. Original data downloaded from ABS ??????????????

[https://drive.google.com/](https://drive.google.com/a/stmonicas-epping.com/folderview?id=0BwscJOIIBPX0ZldhM1ZlMXZ6cDQ&usp=sharing)